

Student growth in Indiana Fundamental principals to guide decisions

Dr. Damian W Betebenner
National Center for the Improvement of Educational Assessment

Overview

- Indiana has been at the forefront nationally in the use of student growth measures.
- Recent policy and legislative changes in Indiana have led to a re-evaluation of the calculation and use of growth.
- In this presentation I hope to do two things:
 - Provide an quick but thorough technical overview of issues related to the calculation and use of student growth.
 - Situate within this overview the current deliberations going on in Indiana.

Growth: Technical Considerations

- It isn't as hard as popular perception suggests!
- Calculation may be complicated, but knowing how to calculate student growth is *not* synonymous with understanding student growth:
 - The number π is difficult to calculate but is simple to understand conceptually.
 - Height and weight percentiles associated with children are difficult to calculate but are simple to understand conceptually.
- Computational details are important, but often obscure a simple conceptual basis.

Growth: Technical Considerations

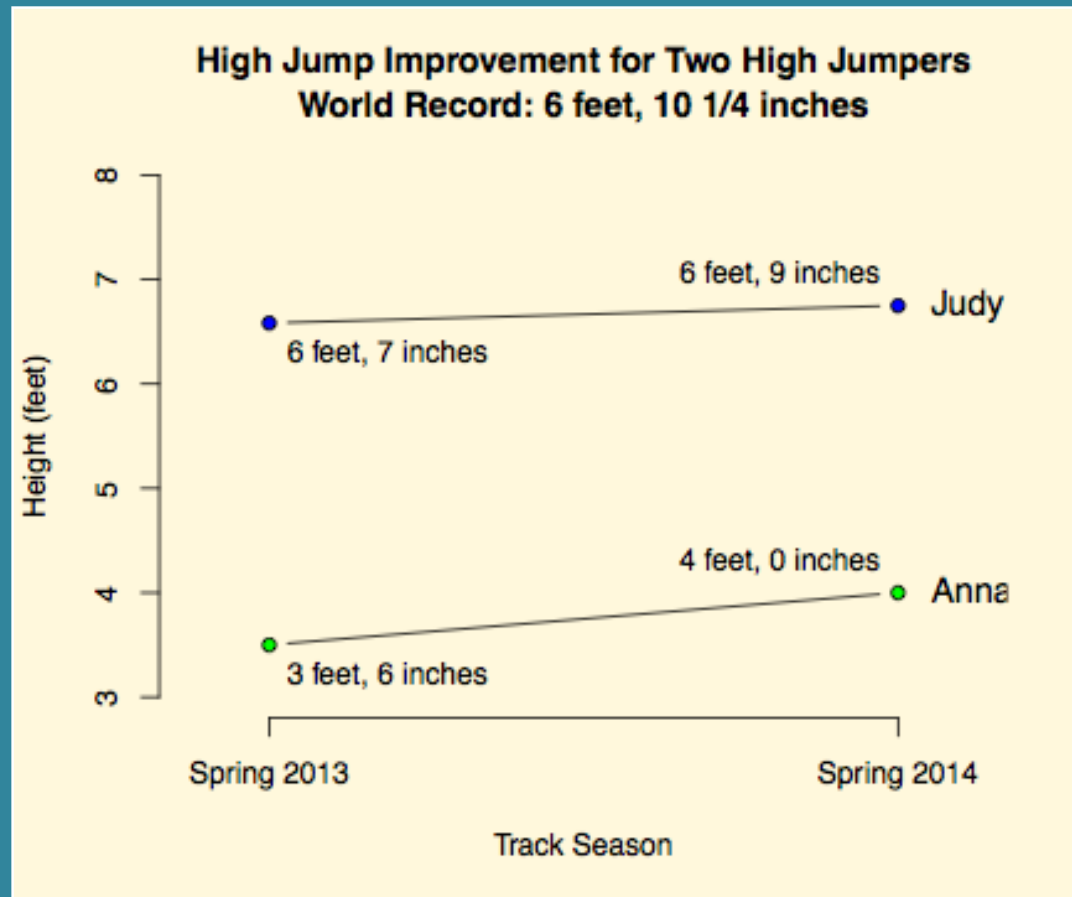
- Indiana's use of student growth has always started with an examination of individual student progress.
- To that end, I will discuss 3 types of growth:
 - Growth as an amount of change (i.e., gain).
 - Growth as a norm/peer referenced quantity.
 - Growth as a criterion referenced quantity.

Growth: Amount of change

- Growth as an amount of change:
 - What could be more simple?
 - Like height or weight just calculate the amount of change.
 - Requires a vertical scale (which Indiana has up to this year and will have going forward) or a set of ordered performance levels.
 - $\text{Growth} = \text{Score Time 2} - \text{Score Time 1}$
- Not as simple as it seems: Comparisons between students often lead to wrong conclusions.

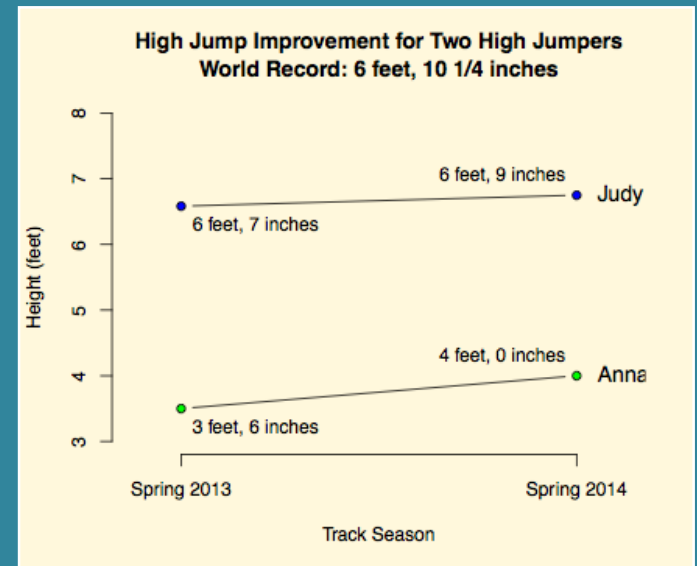
Growth: Amount of change

- Let's step outside education for an example:



Growth: Amount of change

- Which high jumper improved more?
 - Judy improved 2 inches.
 - Anna improved 6 inches.
 - An inch is an inch!!!
 - But is an inch really an inch???
 - 2 inches for an elite jumper is far more remarkable than 6 inches for a novice.



Growth: Amount of change

- The same holds true with the vertically scaled ISTEP+:

	Average Scale Score Change 2013 to 2014
<i>Top 10% ELA Grade 4 2013</i>	-38.8
<i>Bottom 10% ELA Grade 4 2013</i>	+48.5
<i>Top 10% Math Grade 4 2013</i>	-20.3
<i>Bottom 10% Math Grade 4 2013</i>	+54.7

- This isn't meant to impugn the ISTEP+!
- Just as an inch isn't necessary an inch; a scale score point isn't necessarily a scale score point.

Growth: Technical Considerations

- The same issues arise when looking at categorical growth amounts across performance levels:
 - A one category increase from Did Not Pass to Pass is not necessarily equivalent to a one category increase from Pass to Pass+.
 - Growth amounts are ~~tricky~~ to interpret and require *context*.
 - Peer based comparisons used by growth/value-added (categorical or scale score based) analyses are the way to acquire that *context*.

Growth: Peer-based

- Peer-based analyses are also called norm-referenced analyses.
- Like with infant height/weight charts, peers are children/students with similar characteristics.
- For categorical/value-table models, peers are children starting at the same achievement level.
- For scale score based analyses (e.g., SGP, value-added), peers are students with (at least) the same scale score history.

Growth: Technical Considerations

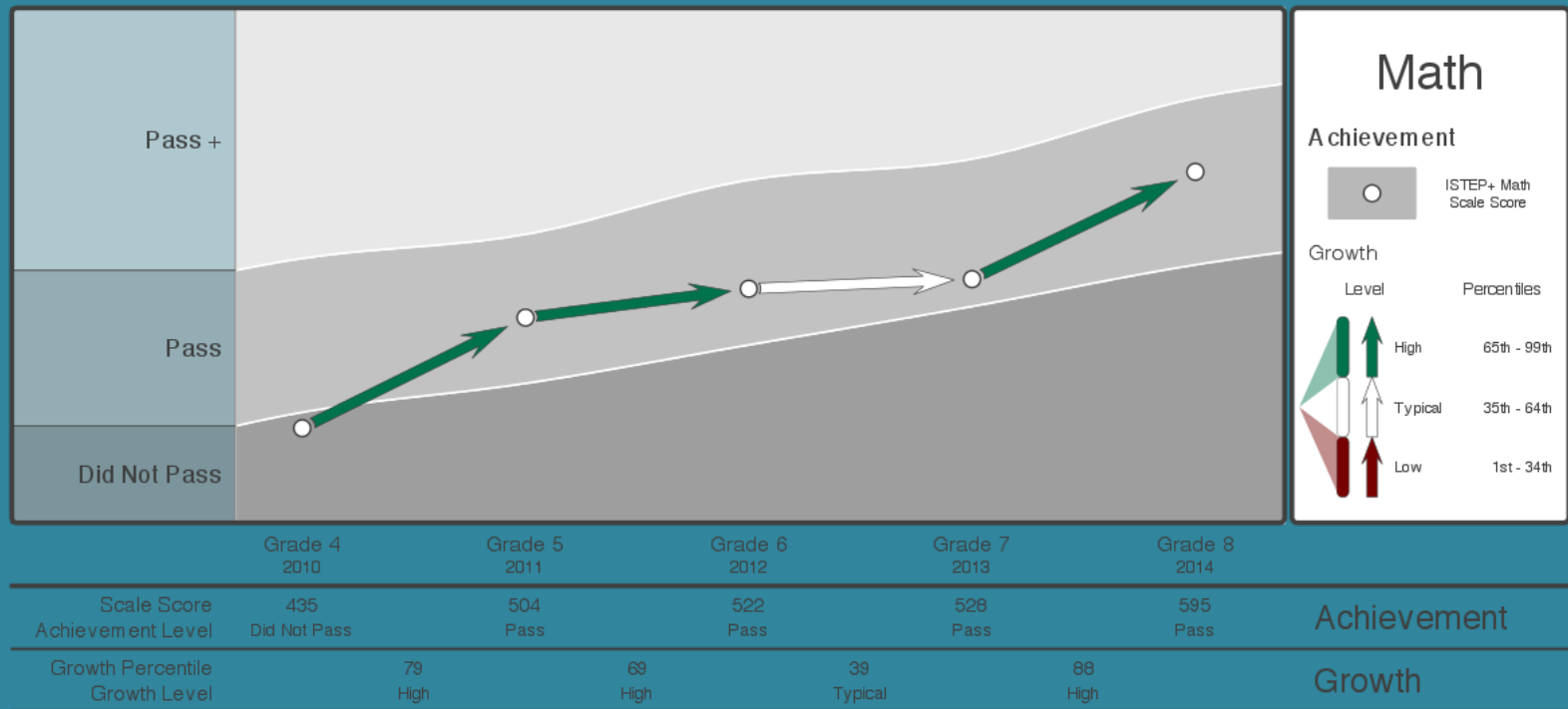
- Categorical/value-tables differentially weight (via impact analyses) categorical transitions to accommodate peer based growth differences.
- Regression based approaches like SGP or value-added utilize analyses based on entire cohorts of students to establish norms against which student progress is examined.
- Peer based norms are invaluable in providing context for stakeholders to understand student growth, whether categorical or scale score based.

Growth: Technical Considerations

- Peer based growth calculations are only as good as the assessments on which they are based.
- With quality assessments, peer based analyses can allow for all students to demonstrate growth.
- However, ceiling or floor effects on a test can impede the analysis of growth for students at the extremes.
- The goal is to be able to demonstrate for stakeholders student growth that is exemplary as well as that which is problematic.

Growth: Technical Considerations

- Achievement over time is growth
- Arrows are baseline referenced growth



Growth: Technical Considerations

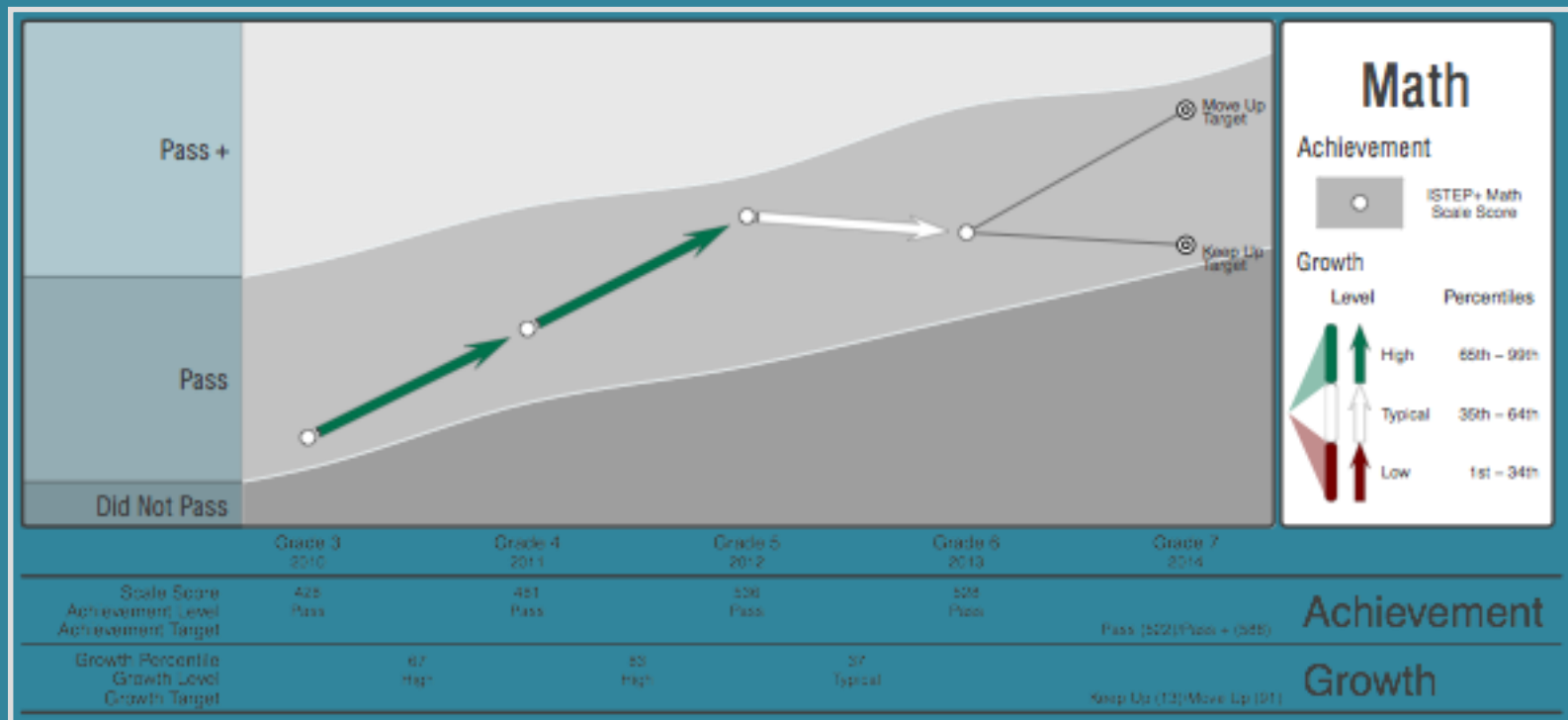
- Peer based comparisons (i.e., norms) are helpful but by their very nature limited.
- One needs criteria as well as norms to enable a complete interpretation.
- A criterion-referenced interpretation examines the quantity against a set of standards.
- Student performance on the ISTEP+ is a familiar example of a criterion referenced interpretation of attainment.

Growth: Technical Considerations

- Criterion-referenced growth is also referred to as growth-to-standard.
- The standards used to examine growth are often the state's performance standards.
- Peer based analyses (e.g., SGP) are used in many states as part of growth-to-standard analyses.
- Growth-to-standard analyses produce growth targets determining whether a student's growth is sufficient to reach/maintain state standards.
- Absent norms, criterion-referenced growth often correlates highly with achievement.

Growth: Technical Considerations

- Achievement over time with growth and growth targets



Growth: Assessment Transition

- The transition from ISTEP+ to the new CCR assessment impacts growth calculations in different ways:
 - It has little to no impact on peer-based analyses.
 - It has a larger impact on criterion-referenced analyses.

Growth: Assessment Transition

<i>Year</i>	<i>Assessment</i>	<i>Methodology</i>	<i>SGP Type</i>
<i>2013-2014</i>	<i>ISTEP+</i>	<i>Current A-F: 1-Yr Growth Targets, updated annually</i>	<i>Peer & criterion growth</i>
<i>2014-2015</i>	<i>Transition to CCR</i>	<i>Current A-F: 1-Yr Growth Targets, equi-percentile concordance analysis</i>	<i>Peer growth</i>
<i>2015-2016</i>	<i>CCR</i>	<i>Transition to Baseline Matrices</i>	<i>Peer growth & criterion growth</i>
<i>2016-2017</i>	<i>CCR</i>	<i>Transition to Baseline Matrices</i>	<i>Peer growth & criterion growth</i>
<i>2017-2018</i>	<i>CCR</i>	<i>Baseline Matrices</i>	<i>Peer & criterion growth</i>

Legislation prohibiting the use of “peer-based growth” prohibits the use of any defensible growth measures in the coming assessment transition years.

Growth: Assessment Transition

- Categorical models require subdivide of the state achievement levels (currently Does Not Pass/Pass/Pass+) to detect movement.
- For the ISTEP+, the current maximum number of subdivided levels supported is 8.
- Transition from ISTEP+ to CCR will require establishing new subdivided levels which may/may not support 8 total levels.
- Floor effects (large percentages of students at the floor of the distribution) will require fewer subdivisions.

Growth: Assessment Transition

- Indiana is not alone. All other states are similarly impacted by transitions to new assessments.
- States using criterion-referenced growth analyses will utilize peer-based growth until at least two years of data are available for criterion-referenced growth can be resumed.
- Indiana is considering different observed growth metrics for the implementation of Option D
- States are advised to think long term and build efficiently toward their ideal system 4 years post transition (i.e., 2017-2018).

Growth: Assessment Transition

- The Accountability Systems Review Panel's Option D is an elegant attempt to blend both categorical peer and criterion-referenced growth.
- Options A & B which utilized strictly categorical transitions demonstrated undesirably high correlations with achievement.
- Option C which utilized a variation of peer-based growth was not as elegant as Option D and had slightly less desirable technical qualities.

Growth: Assessment Transition

- Option D had the most desirable technical characteristics and allows Indiana to:
 - Provide both peer and criterion-referenced student growth data to stakeholders.
 - Employ either strict peer or baseline-referenced student growth analyses.
 - Is consistent with the other uses of growth in the state for teacher evaluation, charter school authorization, ... supporting coherent inferences across distinct system levels.

Observed Growth: Baseline-SGPs

Accountability System Review Panel's Recommendation: "Option D"

	Observed Growth					
	Negative Movement		Static Movement		Positive Movement	
Prior Year Status	Target Range	Points	Target Range	Points	Target Range	Points
PP2	1-41	75	42-66	125	67-99	150
PP1	1-39	75	40-64	125	65-99	150
P3	1-36	50	37-61	100	62-99	125
P2	1-34	50	35-59	100	60-99	125
P1	1-31	50	32-56	100	57-99	125
DNP3	1-29	0	30-54	50	55-99	100
DNP2	1-26	0	27-51	50	52-99	100
DNP1	1-24	0	25-49	50	50-99	100

Thank You

Questions?

Thank you

Center for Assessment
www.nciea.org



Dr. Damian Betebenner
dbetebenner@nciea.org